



University of Groningen

## Measuring the nonlinear biasing function from a galaxy redshift survey

Sigad, Y; Branchini, E; Dekel, A

*Published in:*  
Astrophysical Journal

*DOI:*  
[10.1086/309331](https://doi.org/10.1086/309331)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2000

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Sigad, Y., Branchini, E., & Dekel, A. (2000). Measuring the nonlinear biasing function from a galaxy redshift survey. *Astrophysical Journal*, 540(1), 62-73. <https://doi.org/10.1086/309331>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Measuring the Nonlinear Biasing Function from a Galaxy Redshift Survey

Yair Sigad<sup>1</sup>, Enzo Branchini<sup>2</sup>, & Avishai Dekel<sup>1</sup>

## ABSTRACT

We present a simple method for evaluating the nonlinear biasing function of galaxies from a redshift survey. The nonlinear biasing is characterized by the conditional mean of the galaxy density fluctuation given the underlying mass density fluctuation  $\langle \delta_g | \delta \rangle$ , or by the associated parameters of mean biasing  $\tilde{b}$  and nonlinearity  $\tilde{b}$  (following Dekel & Lahav 1999). Using the distribution of galaxies in cosmological simulations, at smoothing of a few Mpc, we find that  $\langle \delta_g | \delta \rangle$  can be recovered to a good accuracy from the cumulative distribution functions of galaxies and mass,  $C_g(\delta_g)$  and  $C(\delta)$ , despite the biasing scatter. Then, using a suite of simulations of different cosmological models, we demonstrate that  $C(\delta)$  is robust compared to the difference between  $C_g$  and  $C$ , and can be approximated for our purpose by a cumulative log-normal distribution of  $1 + \delta$  with a single parameter  $\sigma$ . Finally, we show how the nonlinear biasing function can be obtained with adequate accuracy directly from the observed  $C_g$  in redshift space. Thus, the biasing function can be obtained from counts in cells once the rms mass fluctuation at the appropriate scale is assumed a priori. The relative biasing function between different galaxy types is measurable in a similar way. The main source of error is sparse sampling, which requires that the mean galaxy separation be smaller than the smoothing scale. Once applied to redshift surveys such as PSCz, 2dF, SDSS, or DEEP, the biasing function can provide valuable constraints on galaxy formation and structure evolution.

*Subject headings:* cosmology: theory — cosmology: observation — dark matter — galaxies: distances and redshifts — galaxies: formation — galaxies: clustering — large-scale structure of universe

---

<sup>1</sup>Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

<sup>2</sup>Kapteyn Institute, University of Groningen, Landleven 12, 9700 AV, Groningen, The Netherlands

## 1. INTRODUCTION

The fact that galaxies of different types cluster differently (e.g., Dressler 1980; Lahav, Nemiroff & Piran 1990; Santiago & Strauss 1992; Loveday *et al.* 1995; Hermit *et al.* 1996; Guzzo *et al.* 1997) indicates that the galaxy distribution is in general biased compared to the underlying mass distribution. Cosmological simulations confirm that halos and galaxies must be biased (e.g., Cen & Ostriker 1992; Kauffmann, Nusser & Steinmetz 1997; Blanton *et al.* 1999; Somerville *et al.* 2000). The biasing becomes even more pronounced at high redshift, as predicted by theory (e.g., Kaiser 1986; Davis *et al.* 1985; Bardeen *et al.* 1986; Dekel & Rees 1987; Mo & White 1996; Bagla 1998; Jing & Suto 1998; Wechsler *et al.* 1998), and confirmed by the strong clustering of galaxies observed at  $z \sim 3$  (Steidel *et al.* 1996; 1998). Knowing the biasing scheme is crucial for extracting dynamical information and cosmological constants from the observed galaxy distribution, and may also be very useful for understanding the process and history of galaxy formation.

The simplest possible biasing model relating the density fluctuation fields of matter and galaxies,  $\delta$  and  $\delta_g$ , is the deterministic and linear relation,  $\delta_g(\mathbf{x}) = b\delta(\mathbf{x})$ , where  $b$  is a constant linear biasing parameter. However, this is at best a crude approximation, because it is not self-consistent (e.g., it does not prevent  $\delta_g$  from becoming smaller than  $-1$  when  $b > 1$ ) and is not preserved in time. At any given time, scale and galaxy type, the biasing is expected in general to be nonlinear, i.e.,  $b$  should vary as a function of  $\delta$ . The nonlinearity of dark-matter halo biasing (as well as its dependence on scale, mass and time) is approximated fairly well by the model of Mo & White (1996), based on the extended Press-Schechter formalism (Bond *et al.* 1991). Improved approximations have been proposed by Jing (1998), Sheth & Tormen (1999) and Porciani *et al.* (1999). It is quantified further for halos and galaxies using cosmological  $N$ -body simulations with semi-analytic galaxy formation (e.g., Somerville *et al.* 2000). The biasing is also expected, in general, to be stochastic, in the sense that a range of values of  $\delta_g$  is possible for any given value of  $\delta$ . For example, if the biasing is nonlinear on one scale, it should be different and non-deterministic on any other scale. The origin of the scatter is shot noise as well as the influence of physical quantities other than mass density (e.g., velocity dispersion, dimensionality of local deformation, etc.) on the efficiency of galaxy formation.

Dekel & Lahav (1999) have proposed a general formalism for galaxy biasing, that separates nonlinearity and stochasticity in a natural way. The density fields are treated as random fields, and the biasing is fully characterized by the conditional probability distribution function  $P(\delta_g|\delta)$ . The constant linear biasing factor  $b$  is replaced by a mean *biasing function*,

$$\langle \delta_g | \delta \rangle \equiv b(\delta) \delta, \quad (1)$$

which can in principle take a wide range of functional forms, restricted by definition to have  $\langle \delta_g \rangle = 0$  and  $\langle \delta_g | \delta \rangle \geq -1$  for any  $\delta$ . The stochasticity is expressed by the higher moments about this mean, such as the conditional variance

$$\sigma_b^2(\delta) \equiv \langle \epsilon^2 | \delta \rangle / \sigma^2, \quad \epsilon \equiv \delta_g - \langle \delta_g | \delta \rangle, \quad (2)$$

scaled for convenience by the variance of mass fluctuations,  $\sigma^2 \equiv \langle \delta^2 \rangle$ . To second order, the biasing function  $b(\delta)$  can be characterized by two parameters: the moments  $\hat{b}$  and  $\tilde{b}$ ,

$$\hat{b} \equiv \langle b(\delta) \delta^2 \rangle / \sigma^2 \quad \text{and} \quad \tilde{b}^2 \equiv \langle b^2(\delta) \delta^2 \rangle / \sigma^2. \quad (3)$$

The parameter  $\hat{b}$  is the natural extension of the linear biasing parameter, measuring the slope of the linear regression of  $\delta_g$  on  $\delta$ , and  $\tilde{b}/\hat{b}$  is a useful measure of non-linearity. The stochasticity is characterized independently by a third parameter,  $\sigma_b^2 \equiv \langle \epsilon^2 \rangle / \sigma^2$ . As has been partly explored by Dekel & Lahav (1999), these parameters should enter any nonlinear analysis aimed at extracting the cosmological density parameter  $\Omega$  from a galaxy redshift survey, and are therefore important to measure.

In this paper we propose a simple method to measure the biasing function  $b(\delta)$  and the associated parameters  $\hat{b}$  and  $\tilde{b}$  from observed data that are either already available, such as the PSCz redshift survey (Saunders *et al.* 2000), or that will soon become available, such as the redshift surveys of 2dF (Colless 1999) and SDSS (e.g., Loveday *et al.* 1998) and high-redshift surveys such as DEEP (Davis & Faber 1999). Alternative methods have been proposed to measure the biasing function, using the cumulant correlators of the observed distribution of galaxies in redshift surveys (Szapudi 1998) or their bispectrum (Matarrese, Verde, Heavens 1997, Verde *et al.* 1998).

We first show in §2, using halos and galaxies in  $N$ -body simulations, that the difference between the cumulative distribution functions (CDFs) of galaxies and mass can be straightforwardly translated into  $\langle \delta_g | \delta \rangle$  despite the scatter in the biasing scheme. Then, in §3, we demonstrate that for our purpose,  $C(\delta)$  is insensitive to the cosmological model and can be approximated robustly by a cumulative log-normal distribution. This means that we do not need to observe  $C(\delta)$ , which is hard to do; we only need to measure  $C_g(\delta_g)$  and, independently, the rms value  $\sigma$  of the mass fluctuations on the same scale. In §4, we slightly modify the method to account for redshift-space distortions, and use mock galaxy catalogs from  $N$ -body simulations to evaluate the associated errors. Finally, in §5, we estimate the errors due to the sparse sampling and finite volume. The method and its applications to existing and future data are discussed in §6.

## 2. BIASING FUNCTION FROM DISTRIBUTION FUNCTIONS

Let  $C_g(\delta_g)$  and  $C(\delta)$  be the cumulative distribution functions of the density fluctuations of galaxies and mass respectively (at a given smoothing window). Had the biasing relation been deterministic and monotonic, it could have been determined straightforwardly from the difference between these CDFs at given percentiles,

$$\delta_g(\delta) = C_g^{-1}[C(\delta)] , \quad (4)$$

where  $C_g^{-1}$  is the inverse function of  $C_g$ .<sup>3</sup> In the presence of scatter in the biasing scheme, strict monotonicity is violated, but it is possible that  $C_g^{-1}[C(\delta)]$  is still a good approximation for  $\langle\delta_g|\delta\rangle$ , which *is* likely to be monotonic. The validity of this approximation is addressed in the present section.

We use two cosmological  $N$ -body simulations in which both halos and galaxies were identified (Kauffmann *et al.* 1999). The cosmological models are  $\tau$ CDM (with  $\Omega_m = 1$  and  $h = 0.5$ ) and  $\Lambda$ CDM (with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $h = 0.7$ ).  $N = 256^3$  particles were simulated in a periodic box of comoving size 85 and 141  $h^{-1}$ Mpc respectively (corresponding to a mass resolution of  $1.0 \cdot 10^{10} h^{-1} M_\odot$  and  $1.4 \cdot 10^{10} h^{-1} M_\odot$ ). The simulations were run using a parallel adaptive P<sup>3</sup>M code kindly made available by the Virgo Consortium (see Jenkins *et al.* 1998) as part of the “GIF” collaboration between the HU Jerusalem and the MPA Munich. The present epoch is defined by a linear rms density fluctuation in a top-hat sphere of radius  $8 h^{-1}$ Mpc of  $\sigma_8 = 0.6$  in the  $\tau$ CDM simulation and  $\sigma_8 = 0.9$  in the  $\Lambda$ CDM simulation. Dark-matter halos were identified at densely sampled time steps using a friends-of-friends algorithm. Galaxies were identified inside these halos by applying in retrospect semi-analytic models (SAMs) of galaxy formation (Kauffmann *et al.* 1999). The SAMs simulate the important physical processes of galaxy formation such as gas cooling, star formation and supernovae feedback. At different times in the redshift range 0 to 3, we select halos by mass and galaxies by luminosity or type. We then compute density fields by applying top-hat smoothing with radii in the range  $5 - 15 h^{-1}$ Mpc. We report detailed results for the case of  $8 h^{-1}$ Mpc smoothing, and refer to the scale dependence in several places.

The figures of this section illustrate the success of the approximation, equation (4), in several different cases based on the  $\tau$ CDM simulation, with top-hat smoothing of radius  $8 h^{-1}$ Mpc (hereafter TH8, or THX for radius  $X h^{-1}$ Mpc), and at different redshifts. Figure 1 refers to halos of mass  $> 10^{12} h^{-1} M_\odot$  ( $> 100$  particles). On the top we show the cumulative distributions of halos and underlying mass fluctuations,  $C_g(\delta_g)$  and  $C(\delta)$  (our notation does not distinguish between halos and galaxies). The errors in  $C_g$  are computed from 20 bootstrap simulations of the halo field. The errors in  $C$ , estimated in the same way, are smaller by an order of magnitude and are therefore not shown. The bottom panels show a point-by-point comparison of the TH8 fields of  $\delta_g(\mathbf{x})$  and  $\delta(\mathbf{x})$  at points randomly chosen (1:8) from a uniform grid of spacing  $2.64 h^{-1}$ Mpc within the simulation box. The true mean biasing function  $\langle\delta_g|\delta\rangle$  is marked by the filled circles with attached error bars. It is computed by a local linear regression of  $\delta_g$  on  $\delta$  within each bin of  $\delta$ , adopting the value of the fitted line at the center of the bin (only every other bin is shown). Shown in comparison (solid line) is the approximation for  $\langle\delta_g|\delta\rangle$  obtained by equation (4) from the CDFs, and the corresponding  $1\sigma$  error range based on the bootstrap realizations (dotted lines).

As can be seen in Figure 1, the approximation is excellent over most of the  $\delta$  range —

---

<sup>3</sup>A similar relation has been used by Narayanan & Weinberg (1998) for “debiasing” the galaxy density field for the purpose of dynamical reconstruction.

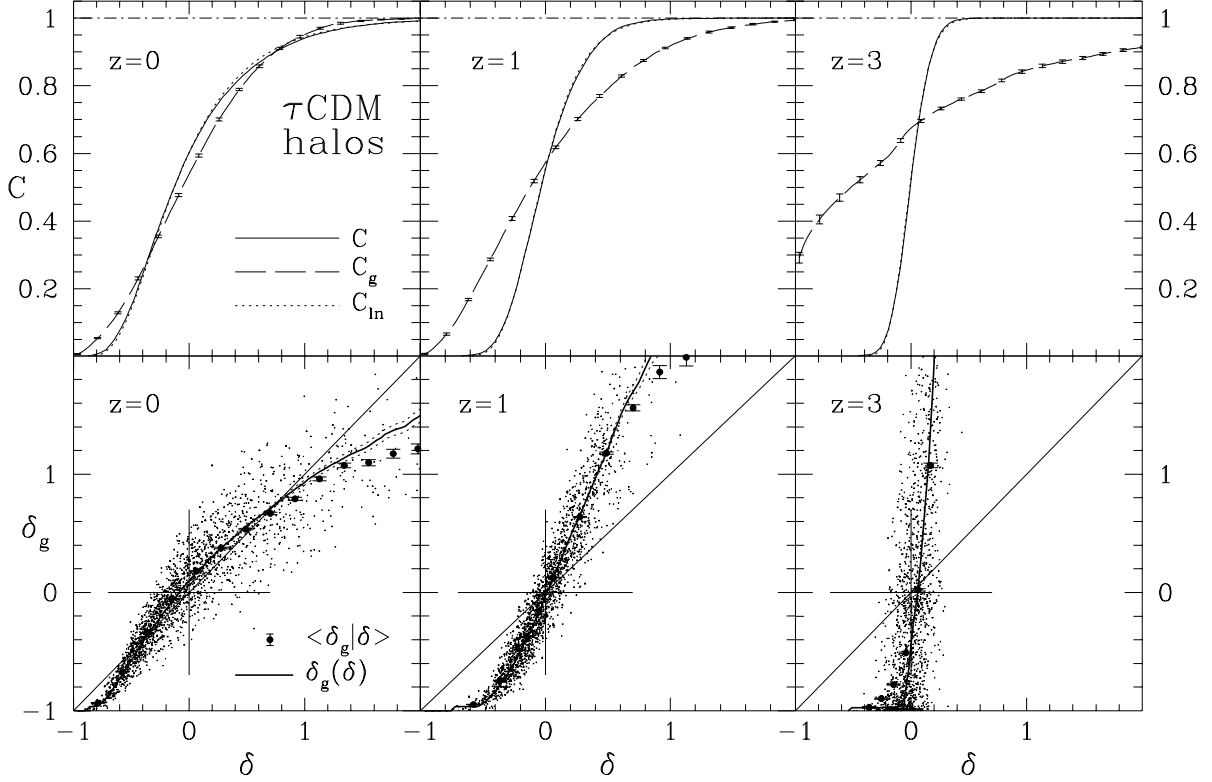


Fig. 1.— CDFs and the biasing function at different redshifts for  $\tau$ CDM halos with  $M > 10^{12}h^{-1}M_{\odot}$  and TH8 smoothing. *Top panels:* the matter  $C(\delta)$  (solid) and the halo  $C_g(\delta_g)$  (dashed). Also shown is a log-normal distribution (dotted), largely hidden behind the exact mass distribution. *Bottom Panels:*  $\delta_g(\mathbf{x})$  versus  $\delta(\mathbf{x})$  at grid points within the simulation box. The true mean biasing function  $\langle \delta_g | \delta \rangle$  is marked by the filled circles with error bars. Shown in comparison (solid line) is the approximation obtained by equation (4) from the CDFs and the corresponding  $1\sigma$  error range (dotted).

the deviation at  $z=0$  is within the  $1\sigma$  errors up to  $\delta \sim 1.4$  (corresponding to  $\sim 97\%$  of the volume). Systematic deviations show up at higher  $\delta$  values, where the scatter becomes larger and the mean biasing function flatter, making the deviations from monotonicity larger. In order to quantify the quality of the approximation, we average the residuals (scaled by  $\sigma_g$ ):

$$\Delta = \frac{1}{N_{\text{bins}}\sigma_g^2} \sum_{\delta \text{--bins}}^{N_{\text{bins}}} [\delta_g(\delta) - \langle \delta_g | \delta \rangle]^2, \quad (5)$$

where  $\delta_g(\delta)$  is obtained via equation (4). We exclude the poorly recovered high-density tail by limiting the summation to those  $N_{\text{bins}}$  bins of  $\delta$  for which  $C(\delta) < 0.99$  and  $C_g(\delta_g) < 0.99$ . The values of  $\Delta$  in the various cases studied, including halos and galaxies in  $\tau$ CDM and  $\Lambda$ CDM at different redshifts, are listed in Table 1. For example, for the halos shown in Figure 1 at  $z = 0$  we obtain  $\Delta = 0.08$ , indicating that the typical error in the approximation

$\delta_g(\delta)$  is small compared to the actual scatter  $\sigma_g$  in the halo density field.

A complementary approach for quantifying the quality of the approximation is by testing how well it recovers the values of the moments of the biasing function,  $\hat{b}$  and  $\tilde{b}$ . In Table 1 we present the values of these moments for the different cases, as computed directly from the simulation and as approximated by  $\delta_g(\delta)$  (denoted by a subscript “a”). These biasing parameters are computed based on 99.9% of the volume, excluding the very highest density peaks, where the error is large (The only exception is at  $z=3$ , where we use only 99% of the volume because the errors are even larger). For the halos shown in Figure 1 at  $z = 0$ , we see that  $\hat{b}$  and  $\tilde{b}$  are recovered with errors of 1% and 3% respectively.

The middle panels of Figure 1 refer to  $z = 1$ , where  $\hat{b} \simeq 2.2$ . The approximation of equation (4) holds well in this case up to  $\delta \sim 0.7$ , which corresponds to  $\sim 98\%$  of the volume. The approximation remains good despite the large scatter (compared to the  $z = 0$  case) because the steepness of the biasing function helps maintaining reasonable monotonicity. The goodness of the recovery of the biasing function, with  $\Delta = 0.07$ , is similar to the  $z = 0$  case. The parameters  $\hat{b}$  and  $\tilde{b}$  are recovered with an accuracy of  $\sim 5\%$  (Table 1). The right panels of Figure 1 demonstrate that the approximation is valid even at  $z=3$ , where

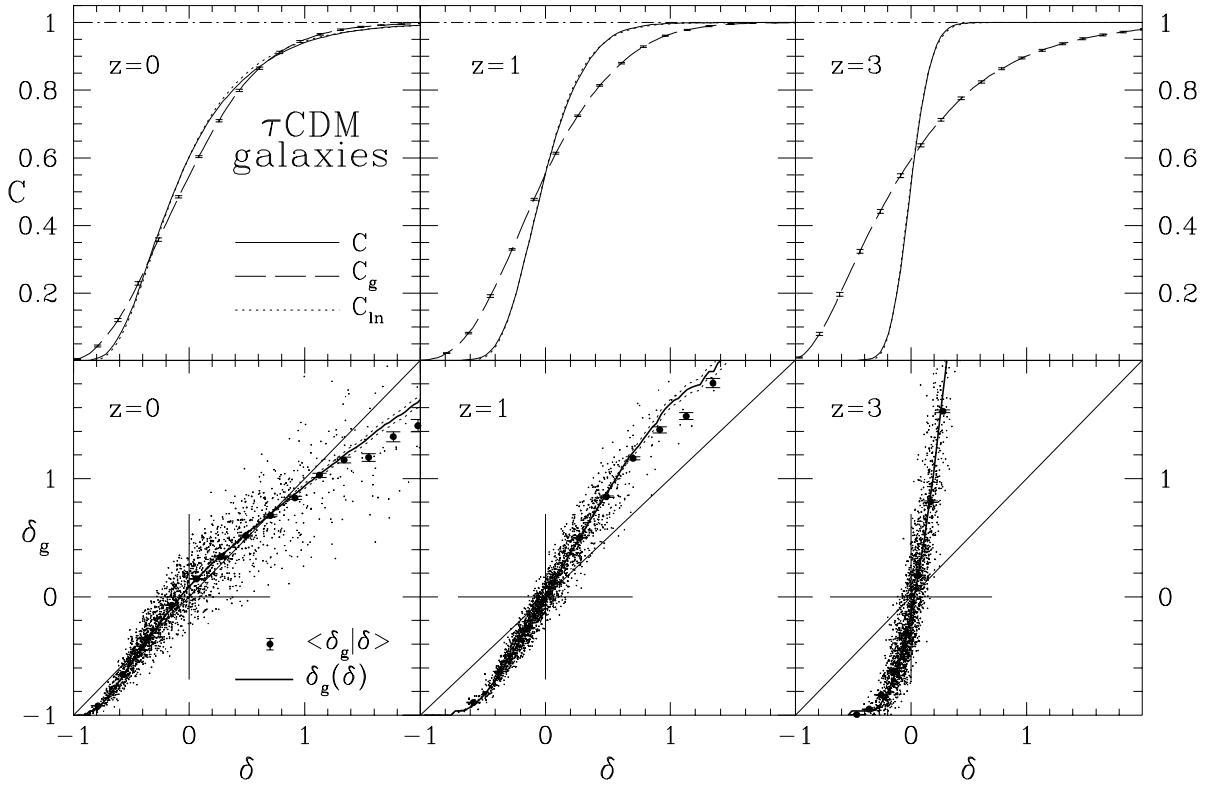


Fig. 2.— Same as Figure 1, but for bright galaxies of  $M_B < -21$  rather than massive halos.

the biasing is extremely strong,  $\hat{b} \simeq 6.6$ . The recovery of the biasing function is still good,  $\Delta = 0.20$ , and its moments are approximated to within  $\sim 2\%$ .

The halo biasing function in the  $\Lambda$ CDM cosmology is recovered, in general, with similar success, as can be seen in the top part of Table 1. Note that in this case the recovery actually improves at higher redshift. This reflects the fact that in  $\Lambda$ CDM the halo biasing scatter becomes smaller at higher redshift (see Somerville *et al.* 2000, Fig. 17). It results from the smaller shot noise due to the higher abundance of high-redshift halos in  $\Lambda$ CDM compared to  $\tau$ CDM.

Figure 2 is analogous to Figure 1, but now for bright galaxies of  $M_B - 5 \log h < -19.5$ . The recovery is again quantified in Table 1; it is quite similar to the case of halos. The typical error is  $\Delta \leq 0.08$ , and the biasing parameters are recovered with an error of a couple to a few percent.

The performance of our method has been tested for smoothing scales in the range  $5 - 15 h^{-1} \text{Mpc}$ . For the  $\tau$ CDM model, we find that the quality of the approximation is practically independent of scale throughout this range; the relative error in the biasing parameters is at the level of a few percent, and  $\Delta$  is in the range 0.1 to 0.2, rather similar to the values quoted in Table 1 for TH8 smoothing. On the other hand, for  $\Lambda$ CDM we do find that the performance improves with increasing smoothing scale. With TH15 at  $z = 0$ , for halos (or galaxies), the errors in the biasing parameters reduce to below 3% (1%), and  $\Delta = 0.07$  (0.04), while for TH5 smoothing these errors are about 4 times larger. This difference between the two models can be attributed to a difference in the scale dependence of the biasing scatter (Somerville *et al.* 2000, Figure 16), which translates to an error in our method via increased deviations from monotonicity.

Before we proceed with the biasing relative to the underlying mass, we note that the *relative* biasing function of two different galaxy types,  $\langle \delta_{g_2} | \delta_{g_1} \rangle$ , can be directly observable from a redshift survey. Again, for a deterministic and monotonic biasing process one has

$$\delta_{g_2}(\delta_{g_1}) = C_{g_2}^{-1}[C_{g_1}(\delta_{g_1})] , \quad (6)$$

and when biasing scatter is present, the question is to what extent equation (6) provides a valid approximation for the true relative biasing function.

Figure 3 shows the relative biasing function of “early” and “late” type galaxies in the two cosmological models, at  $z = 0$  and with TH8 smoothing. These galaxy types are distinguished in the SAM  $N$ -body simulations according to the ratio of bulge to total luminosity in the V band being larger or smaller than 0.4 respectively. The large scatter in the relative biasing, due to errors in the two density fields, is reduced by including all the galaxies, without applying a luminosity cut.

As can be seen in the last three columns of Table 1, the quality of the recovery of the relative biasing function is not as good as in the case of the absolute biasing of galaxies or halos. The values of  $\Delta$  range from 0.2 to 0.56, compared to 0.08 to 0.16 in the former cases. This is expected, because in the case of relative biasing the two density fields contribute to



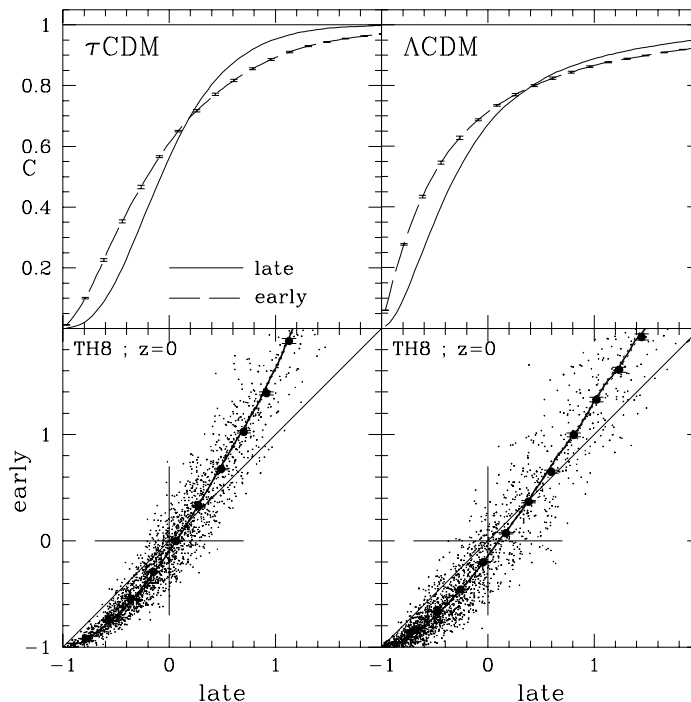


Fig. 3.— The relative biasing of early versus late type galaxies, at  $z=0$ , for  $\tau$ CDM (right panels) and  $\Lambda$ CDM (left panels). The symbols are as in Figure 1.

the stochasticity or deviation from monotonicity (see also the important role of sampling errors in the recovery of the biasing function, §4.2). The moments of the relative biasing function are recovered with better than 15% accuracy at  $z \leq 1$ , and to  $\sim 25\%$  accuracy at  $z = 3$ , in both cosmologies. In calculating the moments, unlike in Figure 3, a luminosity cut has been applied:  $M_B - 5 \log h < -19.5$ , and 99% of the volume was used. The fact that the  $\Delta$  values are still significantly smaller than unity and the errors in the biasing parameters are not larger than 25% indicate that our method is capable of yielding meaningful estimates of the relative biasing function. In both cosmologies, the relative biasing is almost scale independent in the range  $5 - 15 \ h^{-1}\text{Mpc}$ , as is the quality of the reconstruction.

### 3. THE MASS CDF: ROBUST AND LOGNORMAL

Large redshift surveys provide a rich body of data for mapping the galaxy density field in extended regions of space and computing its CDF with adequate accuracy. However, direct mapping of the *mass* density field is much harder. For example, POTENT reconstruction from peculiar velocities (Dekel, Bertschinger & Faber 1990; Dekel *et al.* 1999; Dekel 2000) yields the mass distribution in our local cosmological neighborhood (even out to  $\sim 100 \ h^{-1}\text{Mpc}$ ), which in principle enables direct mapping of the local biasing field. However, the sparse and noisy data limit the mass reconstruction to low resolution ( $\sim 10 \ h^{-1}\text{Mpc}$ )

compared to the volume sampled, which introduces large cosmic scatter in the mass CDF. New accurate data nearby, based on SBF distances (Tonry *et al.* 1997) do enable a promising resolution of a few Mpc (see Dekel 2000), but limited to inside the local sphere of radius  $\sim 30 h^{-1}\text{Mpc}$ .

What makes the method proposed here feasible is the fact that the mass CDF is only weakly sensitive to variations in the cosmological scenario within the range of models that are currently considered as viable models for the formation of large-scale structure (e.g., Primack 1998, Bahcall *et al.* 1999). It has been proposed that the mass PDF can be well approximated by a log-normal distribution in  $\rho/\bar{\rho} = 1 + \delta$  (e.g., Coles & Jones 1991; Kofman *et al.* 1994), and it has since been argued that this approximation becomes poor for certain power spectra and at the tails of the distribution (Bernardeau 1994; Bernardeau & Kofman 1995). In this section, we investigate the robustness of  $C(\delta)$  for our purpose here, namely, in comparison with the typical difference between the CDFs of galaxies and mass (i.e., the mean biasing function) which we are trying to approximate.

We use for this purpose a suite of  $N$ -body simulations of six different cosmological models. In addition to the two high-resolution simulations of  $\tau\text{CDM}$  and  $\Lambda\text{CDM}$  used in the previous section, we have simulated three random realizations of each of the three following models (all using a Hubble constant of  $h = 0.5$ ): standard CDM (SCDM;  $\Omega_m = 1$  with spectral index  $n = 1$ ), an extreme open CDM (OCDM;  $\Omega_m = 0.2$ ,  $n = 1$ ), and an extreme tilted CDM (TCDM;  $\Omega_m = 1$ ,  $n = 0.6$ ). These simulations were run by Ganan *et al.* (2000, in preparation) using a PM code (by Bertschinger & Gelb 1991), with  $128^3$  particles in a  $256 h^{-1}\text{Mpc}$  box. The present epoch is defined in these simulations by a linear fluctuation amplitude of  $\sigma_8 = 1.0$ . A similar simulation was run using a constraint realization (CR) of the local universe based on the galaxy density in the IRAS 1.2Jy redshift survey under the assumption of no biasing (Kolatt *et al.* 1996), with  $\Omega_m = 1$  and the present defined in this case by  $\sigma_8 = 0.7$ .

Figure 4 (left) shows for the different models the deviations  $\Delta C(\delta)$  of the mass CDFs, smoothed TH8, from a cumulative log-normal distribution with the same  $\sigma$ . The log-normal probability density is

$$P_{\ln}(\delta) = \frac{1}{\tilde{\rho}} \frac{1}{\sqrt{2\pi}s} \exp \left[ -\frac{(\ln \tilde{\rho} - m)^2}{2s^2} \right] , \quad (7)$$

where

$$\tilde{\rho} = 1 + \delta , \quad m = -0.5 \ln(1 + \sigma^2) , \quad s^2 = \ln(1 + \sigma^2) \quad \text{and} \quad \sigma^2 = \langle \delta^2 \rangle . \quad (8)$$

The cumulative log-normal distribution is obtained by integration,

$$C_{\ln}(\delta) = \text{erf} \left[ \frac{\ln \tilde{\rho} - m}{s} \right] , \quad (9)$$

where

$$\text{erf}(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt . \quad (10)$$

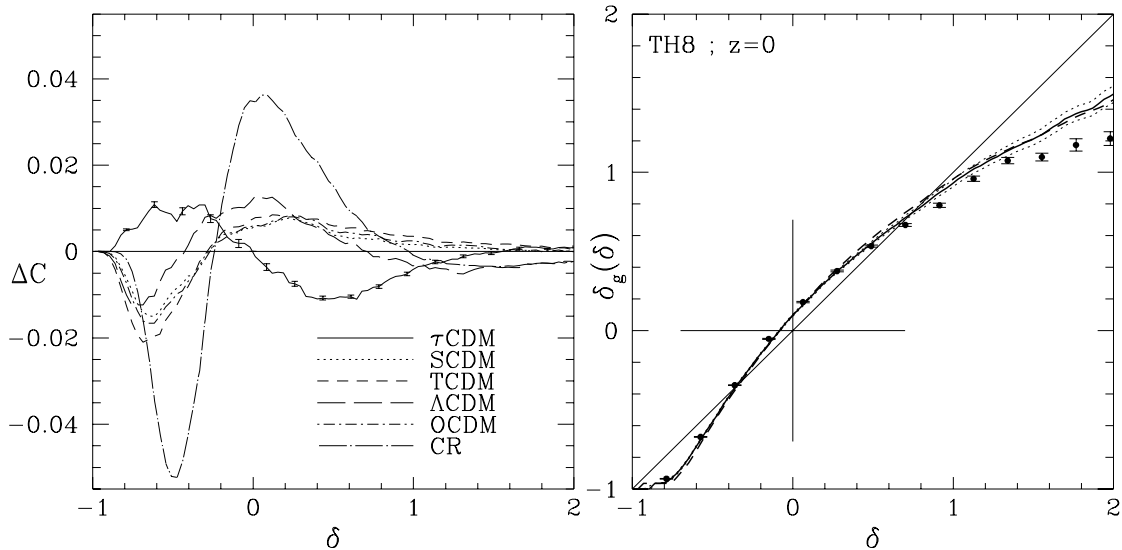


Fig. 4.— Robustness of the mass CDF to cosmological models. *Left*: The deviation  $\Delta C$  of the CDFs from a cumulative log-normal distribution, for various CDM cosmologies at  $z = 0$ :  $\tau$ CDM (solid);  $\Lambda$ CDM (long-dashed); OCDM (dot-dashed); TCDM (dashed); SCDM (dotted); and CR (dot-long-dashed). *Right*: The approximation  $\delta_g(\delta)$  based on the exact  $C(\delta)$  (solid curve, with dotted lines marking 1- $\sigma$  errors), versus the one based on the approximation  $C(\delta) = C_{\ln}(\delta)$  instead (dashed curve). They lie almost on top of each other. The true mean biasing function  $\langle \delta_g | \delta \rangle$  is shown for comparison (points with error bars). All are for halos with  $M > 10^{12} h^{-1} M_\odot$  in the  $\tau$ CDM simulation.

For the cases of OCDM, TCDM and SCDM, the CDF is obtained from the three simulations of each model put together. The errors are similar in the different cases; we therefore plot representative error bars only for the  $\tau$ CDM case.

In all the realizations that had random Gaussian initial conditions, the deviation from lognormality is less than 2%. The constrained realization shows somewhat larger deviations, but even in this case they never exceed 5%. These deviations are indeed smaller than the typical differences between  $C_g(\delta)$  and  $C(\delta)$ , which are on the order of 10% (see Figure 1).

In order to evaluate how important the contribution of  $\Delta C$  is to the error in the recovery of  $\langle \delta_g | \delta \rangle$ , we compare in the right panel of Figure 4 the true  $\langle \delta_g | \delta \rangle$  in the  $\tau$ CDM simulation with two approximations  $\delta_g(\delta)$  based on equation (4), one using the true matter CDF and the other replacing it with a cumulative log-normal distribution of the same  $\sigma$ . The results of the two approximations are very similar; the differences between them seem to be much smaller than the differences between each of them and the true biasing function  $\langle \delta_g | \delta \rangle$ . We can conclude that for the purpose of recovering the biasing function, for the range of Gaussian cosmological models considered,  $C_{\ln}$  is a good approximation for  $C$ .

The proximity of  $C$  and  $C_{\ln}$  could have been alternatively evaluated by the Kolmogorov-Smirnov (KS) statistic,  $D = \max\{|\Delta C|\}$ . For computing the KS significance  $q(D)$ , we

estimate the effective number of “independent” points by  $N_{\text{eff}} = V_{\text{box}}/V_{\text{win}}$ , where  $V_{\text{box}}$  is the volume of the simulation box and  $V_{\text{win}}$  is the effective volume of the smoothing window. A value of  $q \simeq 1$  ( $D \ll 1$ ) corresponds to a good match, and  $q \ll 1$  ( $D \simeq 1$ ) to a poor match. For our  $\tau$ CDM simulation, with TH8 smoothing at  $z = 0$  and 1, we obtain  $D \simeq 0.01$  and  $q > 0.9999$ , confirming that  $C_{\text{ln}}$  is a good fit. However, for the larger SCDM and OCDM simulations, although  $D$  is still only  $\simeq 0.015$ , the corresponding  $q$  values are at the level of only a few percent. For TCDM and CR, where  $D$  is 0.016 and 0.052 respectively, the values of  $q$  drop to the level of a fraction of a percent, and the discrepancy seems large. This KS test indicates that the log-normal approximation is not always perfect for general purpose, as has been argued in the literature. However, our direct tests reported above demonstrate that the use of the log-normal approximation is adequate for the recovery of the mean biasing function in all these cases.

We comment in passing that while the mass CDF is well approximated for our purpose by a log-normal distribution, the shape of the halo (or galaxy) CDF is usually far from a log-normal shape. This is implied by equation (4), from which it follows that  $C_g(\delta_g) = C[\delta_g^{-1}(\delta_g)]$ . One does not expect to recover a log-normal distribution from a general functional form for  $\delta_g^{-1}$ . In particular, the linear biasing model, which seems to be an acceptable approximation in some cases with large smoothing (e.g., IRAS 1.2Jy galaxies at  $12 h^{-1}\text{Mpc}$  Gaussian smoothing; Sigad *et al.* 1998), leads to a  $C_g(\delta_g)$  that is far from log-normal. Trying to evaluate the difference between  $C_g$  and a log-normal distribution using the KS statistic, we obtain for the halos in the  $\tau$ CDM simulation, with TH8 smoothing, both at  $z = 0$  and 1,  $D \simeq 0.08$  and  $q \simeq 0.05$ , namely a poor fit compared to the  $q \simeq 1$  of  $C$  vs  $C_{\text{ln}}$ . Similar conclusions are valid for galaxies.

#### 4. REDSHIFT DISTORTIONS

The densities as measured in redshift space (z-space) are in general different from the real-space (r-space) densities addressed so far, because the radial peculiar velocities distort the volume elements along the lines of sight. One approach to deal with redshift distortions is to start by recovering the full galaxy density field in r-space, using the linear or a mildly-nonlinear approximation to gravitational instability (e.g., Yahil *et al.* 1991; Strauss *et al.* 1992; Fisher *et al.* 1995; Sigad *et al.* 1998), and then compute the biasing function in r-space as outlined above. The accuracy of such a procedure would be limited by the approximation used for nonlinear gravity. Another difficulty with this approach is that it requires one to assume a priori a specific biasing scheme, already in the force calculation that enters the transformation from z-space to r-space, while this biasing scheme is the very unknown we are after; this would require a nontrivial iterative procedure.

The alternative we propose here is to actually use the z-space CDF,  $C_{g,z}(\delta_{g,z})$ , as provided directly from counts in cells of galaxies in a redshift survey. If the redshift distortions affect the densities of galaxies and mass in a similar way, then one expects the biasing function in

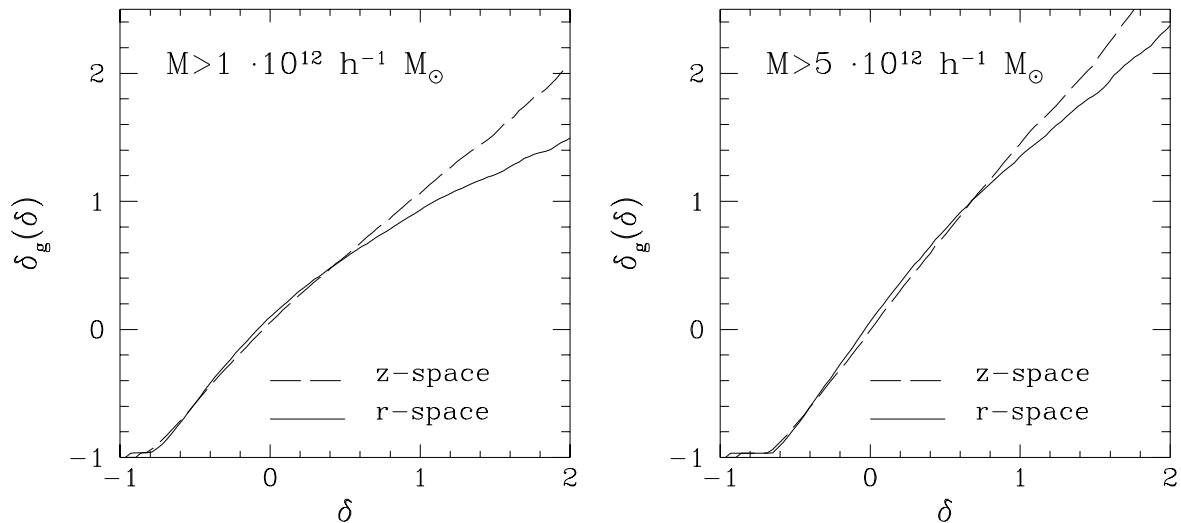


Fig. 5.— Biasing functions in z-space (dashed) versus r-space (solid). The biasing functions are derived from the corresponding TH8 CDFs of halos and mass in the  $\tau$ CDM simulation at  $z = 0$ . Shown are halos of  $M > 10^{12} h^{-1} M_{\odot}$  (left) and  $M > 5 \cdot 10^{12} h^{-1} M_{\odot}$  (right).

z-space to be similar to the one in real space,

$$\langle \delta_{g,z} | \delta_z = \delta \rangle = \langle \delta_g | \delta \rangle . \quad (11)$$

If we only had a robust functional form for the mass CDF in z-space,  $C_z(\delta_z)$ , then we could compute the desired biasing function all in z-space, using equation (4) but with the analogous z-space quantities. We thus need to test the validity of equation (11), and come up with a useful approximation for  $C_z(\delta_z)$ .

Figure 5 illustrates the accuracy of equation (11). It compares the biasing functions in z-space and r-space, as derived via equation (4) and its z-space analog from the corresponding CDFs of halos and mass in the  $\tau$ CDM simulation with TH8 smoothing. The two curves are remarkably similar for  $\delta < 0.6 - 0.8$ , roughly out to the 1-sigma rms fluctuation value. This is roughly the range where the biasing scatter is reasonably small and our basic method is applicable (§2, Figure 1). The curves deviate gradually as  $\delta$  increases due to stronger “fingers of god” effects at high densities. The deviation is somewhat weaker for larger-mass halos (perhaps due to a lower velocity dispersion for more massive objects).

Note that while the results of Figure 5 based on our high-resolution  $\tau$ CDM simulation are quite accurate in the way they treat halos, they may suffer from significant cosmic variance due to the relatively small volume sampled, where the presence (or absence) of a few “fingers of god” could strongly affect the biasing function in the high- $\delta$  regime. To test the validity of equation (11) with reduced cosmic variance, we appeal to yet another set of N-body simulations (by Cole *et al.* 1997) which cover a much larger volume but with lower

resolution. These simulations followed the evolution of  $N = 192^3$  particles in a periodic box of comoving side  $L = 345.6 h^{-1}\text{Mpc}$  using an Adaptive P<sup>3</sup>M code. The cosmological models are  $\Lambda\text{CDM}$  ( $\Omega_{\text{m}} = 0.3$ ,  $\Omega_{\Lambda} = 0.7$ ,  $h = 0.65$ , cluster-normalized to  $\sigma_8 = 1.05$ ) and  $\tau\text{CDM}$  ( $\Omega = 1$ ,  $h = 0.25$ , cluster-normalized to  $\sigma_8 = 0.55$ ). Nine mock catalogs were extracted from each of the parent simulations, each containing  $\sim 5 \cdot 10^5$  particles in a box of  $L = 200 h^{-1}\text{Mpc}$ . The partial overlap between the catalog volumes is thus about 50%. The central “observer” was chosen to mimic certain properties of the Local Group environment (see Branchini *et al.* 1999). Since the resolution of these large simulations is inadequate for a detailed halo identification based on many simulated particles in each halo, we identify individual particles as galaxies using a Monte-Carlo procedure in which the galaxies are chosen to make a random realization of an assumed nonlinear biasing function. Here we adopt the biasing function proposed by Dekel & Lahav (1999) to fit the simulated results of Somerville *et al.* (2000):

$$\delta_{\text{g}}(\delta) = \begin{cases} (1 + b_0)(1 + \delta)^{b_{\text{neg}}} - 1 & \delta \leq 0 \\ b_{\text{pos}}\delta + b_0 & \delta > 0 \end{cases}, \quad (12)$$

with  $b_{\text{neg}} = 2$  and  $b_{\text{pos}} = 1$ . Given the appropriate probability distributions  $P(\delta)$ , the value of  $b_0$  is determined for each choice of the parameters  $b_{\text{neg}}$  and  $b_{\text{pos}}$  such that  $\langle \delta_{\text{g}} \rangle = 0$  as required by definition. We obtain  $b_0 = 0.26$  and  $b_0 = 0.19$  for the models of  $\Lambda\text{CDM}$  and  $\tau\text{CDM}$  respectively.

Figure 6 compares the CDFs and associated biasing functions in r-space and z-space, averaged over nine mock catalog from the large-box  $\Lambda\text{CDM}$  simulation. The z-space biasing function is indeed almost indistinguishable from the r-space one (bottom panels); the differences are typically on the order of a couple of percents. The results for  $\tau\text{CDM}$  are similar.

In order to quantify this difference further, we define a statistic analogous to equation (5):

$$\Delta = \frac{1}{N_{\text{bins}}\sigma_{\text{g}}^2} \sum_{\delta\text{-bins}} [\delta_{\text{g,z}}(\delta_{\text{z}}=\delta) - \delta_{\text{g}}(\delta)]^2, \quad (13)$$

in which the first and second terms are the biasing functions as derived from the CDFs in z-space and r-space respectively. The summation is over bins with  $\delta < \delta_{\text{max}}$ , such that  $\approx 99\%$  of the volume is accounted for. We also compute the two moments of the observed biasing function  $\hat{b}_{\text{obs}}$  and  $\tilde{b}_{\text{obs}}$ . These three quantities, averaged over the mock catalogs, are listed in Table 2 (second column). Their deviation from the “true” values (Table 2, first column) is the systematic error. The quoted errors refer to the  $1\sigma$  scatter about the mean; they represent the random errors. The results are listed for the two models,  $\Lambda\text{CDM}$  and  $\tau\text{CDM}$ . We conclude that the biasing function and its moments, as computed from the z-space CDFs, resemble those computed from the r-space CDFs to within 2%. Note that the Monte Carlo procedure we use to generate mock catalogs artificially reduces the amount of clustering and over-smoothes the density fields for dark and luminous particles. The net effect is to decrease the biasing moments by  $\sim 7\%$ , relative to the values implied by the biasing scheme, equation (12). However, this bias does not affect the present analysis for which “true” values are obtained from the mock catalogs themselves and not from equation (12).

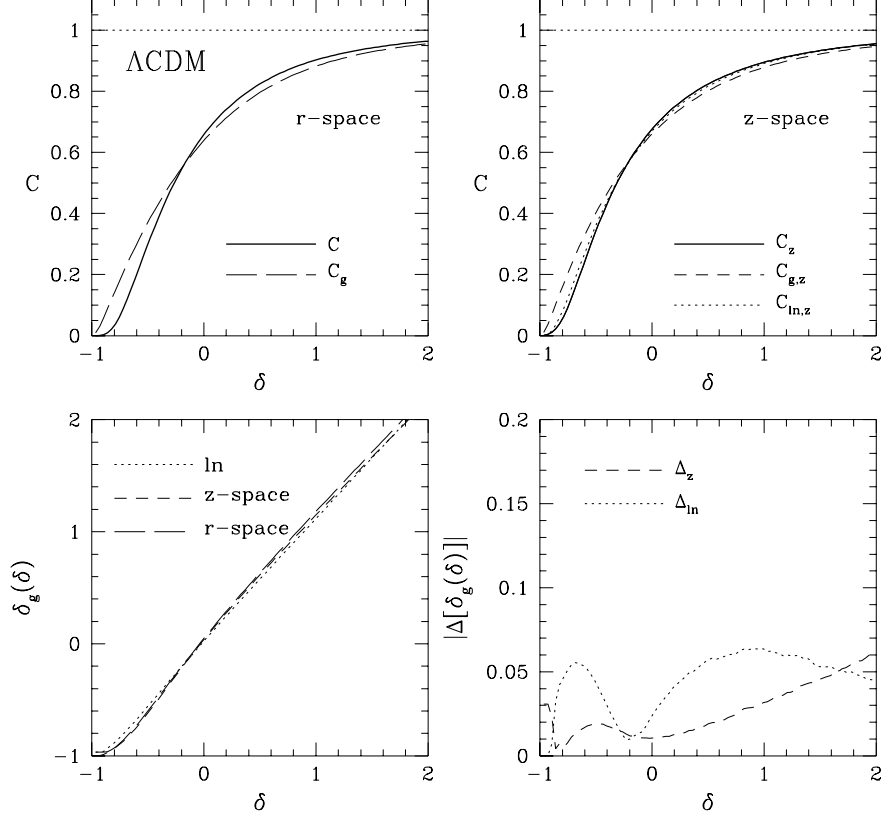


Fig. 6.— CDFs and biasing functions in r-space versus z-space, averaged over mock catalogs that were extracted from the large  $\Lambda$ CDM simulation, with TH8 smoothing. Top: CDFs in r-space (left) and z-space (right), for mass (solid) and for galaxies (dashed). Shown in comparison is  $C_{\ln,z}$  with the  $\sigma_z$  of the matter (dotted). Bottom left: the biasing functions as derived from the CDFs in r-space (long-dash) and z-space (short-dash); they are very similar. Also shown is the biasing function derived in z-space assuming  $C_{\ln,z}$  with  $\sigma_z$  obtained using equation (14) (dotted line). Bottom right: absolute value of the difference between the biasing functions:  $\delta_{g,z}(\delta_z = \delta) - \delta_g(\delta)$  (dashed) and  $\delta_{\ln,z}(\delta_z = \delta) - \delta_g(\delta)$  (dotted).

Our next task is to come up with a robust CDF for the mass in z-space. We try the same log-normal distribution that was found robust for our purpose in r-space (§3), but with a proper rms in z-space,  $\sigma_z$ . Based on the linear approximation for Gaussian fields in the small-angle limit (Kaiser 1987), we express  $\sigma_z$  in terms of  $\sigma$  and  $\Omega_m$  of the cosmological model by:

$$\sigma_z = \left[ 1 + \frac{2}{3}f(\Omega_m) + \frac{1}{5}f^2(\Omega_m) \right]^{1/2} \sigma \quad (14)$$

where  $f(\Omega_m) \simeq \Omega_m^{0.6}$  (with a negligible dependence on  $\Omega_\Lambda$ , see Lahav *et al.* 1991). We thus approximate the z-space biasing function by  $\delta_{\ln,z}(\delta_z)$ , as derived from the z-space CDFs but where the mass CDF is replaced by a cumulative log-normal distribution function  $C_{\ln,z}$  (eq. [9]) with standard deviation  $\sigma_z$  (eq. [14]). The resultant biasing function, averaged over

the mock catalogs, is displayed in the bottom panels of Figure 6. We see that for  $\Lambda$ CDM the differences between  $\delta_{\ln,z}(\delta_z = \delta)$  and  $\delta_g(\delta)$  are at the level of a few percent. For  $\tau$ CDM they are only a bit larger; they exceed 10% but only near  $\delta \sim 2$ , at the tail of the distribution. The error in the biasing function  $\Delta$  defined in analogy to equation (13), and the biasing moments, are listed in Table 2 (third column, marked “z-space ln”). The systematic error  $\Delta$  is still well below 2%, but the biasing parameters are systematically underestimated by 4% and 7% in  $\Lambda$ CDM and  $\tau$ CDM respectively.

Overall, it seems that our straightforward method deals with redshift distortions fairly well, without any *a priori* assumption about the biasing scheme.

## 5. SAMPLING ERRORS

The accuracy of the derivation of the galaxy PDF is limited by two observational factors: the finite volume sampled and the mean density of galaxies in the sample.<sup>4</sup>

In principle, the limited volume is responsible for cosmic variance due to the fact that the long-wavelength fluctuations are not fairly represented in the sampled volume. This is not of major concern for us here because (a) it is expected to introduce only a random error, and (b) as long as the biasing is local, the effects of long waves on the PDFs of galaxies and mass are expected to be correlated, making the local biasing function representative of the universal function despite the relatively small sampling volume.

More important is the *shot noise* introduced by the combination of volume and sampling density effects. For a given cell size (or smoothing length), the error can be divided into the error in the count within each cell and the error due to the finite number of cells in the sample volume. These shot-noise sources may introduce both random and systematic errors. We evaluate them by computing the mean and standard deviation over a suite of mock catalogs in which we vary either the volume or the sampling density for a fixed smoothing scale.

With TH8 smoothing, our mock catalogs from the large  $\Lambda$ CDM simulation contain  $N_{\text{eff}} \sim 3700$  independent cells. However, the currently available redshift surveys allow an analysis in a much smaller volume. For example, a volume-limited subsample from the PSCz catalog (Saunders *et al.* 2000), that is cut at a distance where the average galaxy separation is  $l = 8 h^{-1}\text{Mpc}$  (i.e., on the order of our smoothing scale), contains only  $\sim 600$  independent cells. We therefore estimate the error associated with reducing the sampled volume such that  $N_{\text{eff}} \sim 600$  in each mock catalog. We select from the simulation 9 such non-overlapping sub-volumes, while keeping the sampling density and smoothing length fixed. The results for  $\Lambda$ CDM, averaged over the mock catalogs, are shown in the upper panels of Figure 7, and the results for the two cosmological models are summarized in Table 2 (column 4). We

---

<sup>4</sup>The additional edge effects can be greatly minimized by using a volume-limited sample and a proper choice of cell coverage (see Szapudi & Colombi 1996).



find no significant systematic errors due to the volume effect in a sample like PSCz and with  $\sim 8 h^{-1}\text{Mpc}$  smoothing (except in the very high- $\delta$  tail for  $\tau\text{CDM}$ ). The corresponding random errors in the biasing parameters are 5% and 6% for  $\Lambda\text{CDM}$  and  $\tau\text{CDM}$  respectively.

The sampling density can be parameterized by the mean galaxy separation,  $l$ . In our large simulation  $l = 2.5 h^{-1}\text{Mpc}$ , much smaller than the smoothing length of  $8 h^{-1}\text{Mpc}$ , but in real samples  $l$  could be on the order of the smoothing length. To test the effect of sampling density, we produce 9 mock catalogs in which galaxies are sub-sampled at random from the original catalog such that the mean separation is  $l = 6, 8$ , or  $10 h^{-1}\text{Mpc}$ , while the smoothing length and large volume are kept fixed with  $N_{\text{eff}} \sim 3700$ . The results for  $\Lambda\text{CDM}$  are shown in the bottom panels of Figure 7, and for the two models in Table 2 (columns 5-7). We see that the sparse sampling artificially enhances both positive and negative density fluctuations, which enlarges the width of the galaxy PDF. This results in a steeper biasing function. For  $\Lambda\text{CDM}$ , the effect becomes noticeable only when  $l \geq 8 h^{-1}\text{Mpc}$ , where the

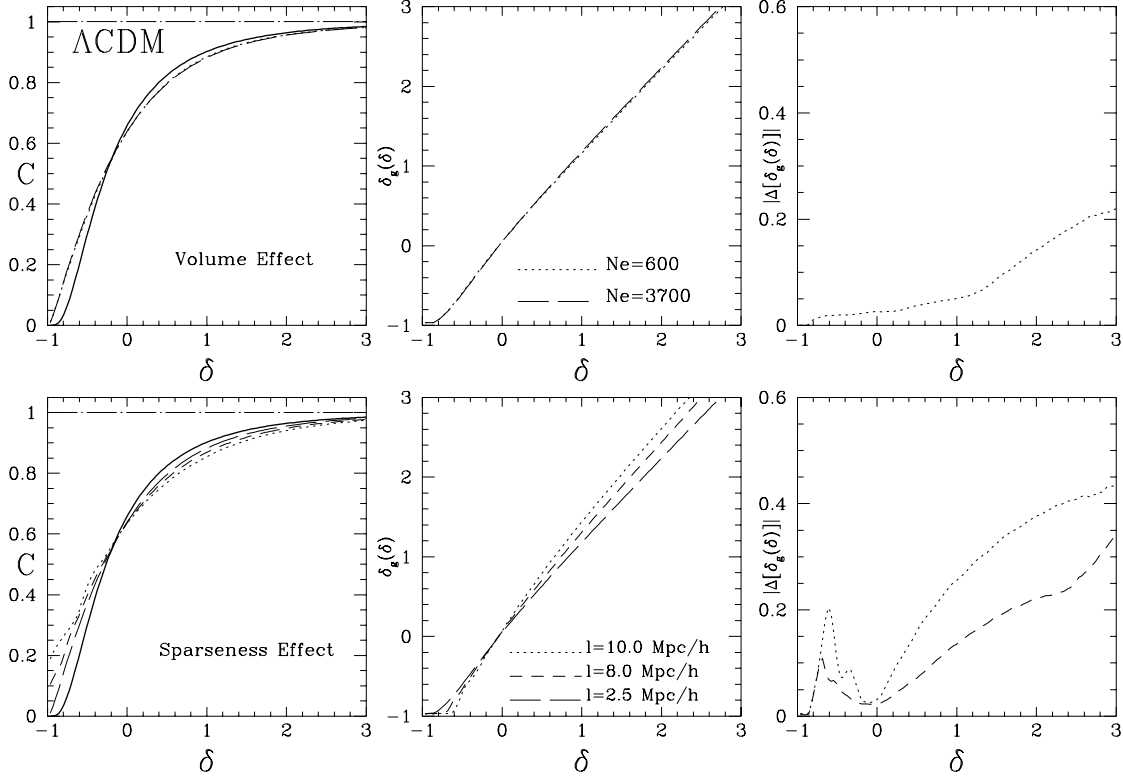


Fig. 7.— Sampling errors due to finite volume (top) and sparse sampling (bottom), for fixed TH8 smoothing, estimated from the large  $\Lambda\text{CDM}$  simulation. Shown are the CDFs in real space (left), the derived biasing function (middle), and the error in it (right). The mass CDF is marked by a solid line, and galaxy CDFs by broken lines. Top: volumes of  $N_{\text{eff}} = 3700$  and  $600$  are marked by long-dashed and dotted lines respectively. Bottom: samples of galaxy separation  $l = 2.5, 8$ , and  $10 h^{-1}\text{Mpc}$  are marked by long-dashed, short-dashed, and dotted lines respectively.

systematic error in the biasing parameters is of order 10% and larger, and  $\Delta$  is of order a few percent. For  $\tau$ CDM the sampling-density effect is noticeable already for  $l \sim 6 h^{-1}\text{Mpc}$ , with the error reaching 30 – 50% at  $l \sim 10 h^{-1}\text{Mpc}$ . A plausible explanation for why the sparse sampling is more damaging in the  $\tau$ CDM model is that the clustering in this model is weaker ( $\sigma_8$  is smaller to match the cluster abundance which constrains  $\sigma_8\Omega^{0.5}$ ), and therefore the high-density regions are poorly sampled by galaxies.

In summary: the main source of error in our analysis is the sparse sampling. For recovering the biasing function with TH8 smoothing, the mean separation should be  $\leq 8 h^{-1}\text{Mpc}$ .

## 6. CONCLUSION

We propose a simple prescription for recovering the mean nonlinear biasing function from a large redshift survey. The biasing function is defined by  $b(\delta)\delta = \langle \delta_g | \delta \rangle$ , and is characterized to second order by two parameters,  $\hat{b}$  and  $\tilde{b}$ , measuring the mean biasing and its nonlinearity respectively. The method is applied at a given cosmology, time, object type and smoothing scale, and involves one parameter that should be assumed a priori — the rms mass density fluctuation  $\sigma$  on the relevant scale.

The main steps of the algorithm are as follows:

1. Obtain the observed cumulative distribution function in redshift space  $C_{g,z}(\delta_{g,z})$ , by counts in cells or with window smoothing at a certain smoothing length.
2. Assume a value for  $\sigma$  on that scale and for the cosmological density parameter  $\Omega_m$ , and approximate the mass CDF in redshift space by  $C_{ln,z}(\delta_z; \sigma_z)$ , the cumulative log-normal distribution (eq. [9]), with the width  $\sigma_z$  derived from  $\sigma$  and  $\Omega_m$  by equation (14).
3. Derive the mean biasing function by

$$\delta_g(\delta = \delta_z) \simeq \delta_{g,z}(\delta_z) = C_{g,z}^{-1}[C_{ln,z}(\delta_z; \sigma_z)] \quad . \quad (15)$$

We first showed that the mean biasing function, at TH8 smoothing, can be derived with reasonable accuracy from the r-space CDFs of galaxies (or halos) and mass, despite the biasing scatter. We then demonstrated that for a wide range of CDM cosmologies the mass CDF can be properly approximated for this purpose by a log-normal distribution of the same width  $\sigma$ . Next we showed that the biasing functions in z-space and r-space are very similar, and that the z-space mass CDF can also be approximated by a log-normal distribution, with a width derived from  $\sigma$  via equation (14). This allows us to apply the method directly to the observed CDF in a redshift survey. The errors in the recovered biasing function and its moments, in an ideal case of dense sampling in a large volume, are at the level of a few percent.

In any realistic galaxy survey the limited volume and discrete sampling introduce further random and systematic errors. For a survey like the PSCz survey, the main source of error is the sampling density; the error does not exceed  $\sim 10\%$  as long as the mean observed galaxy separation is kept smaller than the smoothing radius. We are currently in the process of applying this method to the PSCz survey (E. Branchini, *et al.* 2000, in preparation), where a more specific error analysis will be carried out. The sampling errors are expected to be significantly smaller for the upcoming 2dF and SDSS redshift surveys.

In §2 we showed that our method works well both for halos and for galaxies, on scales 5 to  $15 h^{-1}\text{Mpc}$ , and in the redshift range  $0 \leq z \leq 3$  over which the biasing is expected to change drastically. We obtain a similar accuracy when we vary the cosmological model, the mass of the halos in the comparison, or galaxy properties such as morphological type and luminosity. The approximation  $\delta_g(\delta)$  is consistent (the deviation is less than  $1\text{-}\sigma$ ) with the true average biasing function  $\langle \delta_g | \delta \rangle$  over a wide range of  $\delta$  values, which covers 98 – 99% of the volume, depending on redshift and the type of biased objects. This allows us to estimate the moments of the biasing function to within a few percent (see Table 1). The moments of the biasing function are derived from 99.9% of the volume (99% at  $z=3$  and for relative biasing).

The method requires as external parameters the rms mass-density fluctuation  $\sigma$  and the cosmological parameter  $\Omega_m$ . These can be obtained by joint analyses of constraints from several observational data sets, such as the cluster abundance (e.g., Eke *et al.* 1998), peculiar velocities (e.g., Dekel & Rees 1994; Zaroubi *et al.* 1997; Freudling *et al.* 1999), CMB anisotropies (e.g., de Bernardis *et al.* 1999), and type Ia supernovae (Riess *et al.* 1998; Perlmutter *et al.* 1999). Examples for such joint analyses are Bahcall *et al.* (1999) and Bridle *et al.* (1999).

The method is clearly applicable at  $z \simeq 0$  with available redshift surveys and especially with those that will become available in the near future, 2dF and SDSS. In the future, this method may become applicable at higher redshifts as well, where the biasing plays an even more important role. With the accumulation of Lyman-break galaxies at  $z \sim 3$ , it may soon become feasible to reconstruct their PDF by counts in cells, and our method will allow a recovery of the biasing function at this early epoch, with consequences on galaxy formation and on the evolution of structure.

We have concentrated here on smoothing scales relevant to galaxy biasing, but the method may also be applicable for the biasing of galaxy clusters, on scales of a few tens of Mpc. The biasing scatter may be larger for clusters because of their sparse sampling, but the larger mean biasing parameter for clusters may help in regaining the required monotonicity for equation (4) to provide a valid approximation to the mean biasing function. The mass PDF has been checked to be properly approximated by a log-normal distribution at smoothing scales in the range 20 to  $40 h^{-1}\text{Mpc}$ , using simulations of the standard CDM and Cold+Hot DM models (Borgani *et al.* 1995). The errors due to sparse sampling would require a smoothing scale at the high end of this range.

In a large redshift survey which distinguishes between object types, one can measure the *relative* biasing function between two object types by applying equation (6) in redshift space, using the observed CDFs for the two types without appealing to the underlying mass distribution at all. The upcoming large redshift surveys 2dF and SDSS, and the DEEP survey at  $z \sim 1$ , are indeed expected to provide adequate samples of different galaxy types. Compared with the predictions of simulations and semi-analytic modeling of galaxy formation (e.g., Kauffmann *et al.* 1999; Benson *et al.* 1998; Baugh *et al.* 1999; Somerville & Primack 2000), the measured relative biasing function can provide valuable constraints on the formation of galaxies and the evolution of structure.

While implementing the method outlined above for measuring the mean nonlinear biasing function using current and future redshift surveys, the next challenge is to devise a practical method for measuring the biasing scatter about the mean.

We thank S. Cole, A. Eldar, G. Ganon, T. Kolatt, R. Somerville and our collaborators in the GIF team, J.M. Colberg, A. Diaferio, G. Kauffmann, and S.D.M. White, for providing simulations and mock catalogs. We thank A. Maller, I. Szapudi and D. Weinberg for stimulating discussions. EB thanks the Hebrew University for its hospitality. This work was supported by the Israel Science Foundation grant 546/98.

## REFERENCES

- Bagla, J. S. 1998, MNRAS, 297, 251
- Bahcall, N. A., Ostriker, J. P., Perlmutter, S., & Steinhardt, P. J. 1999, Science, 284, 14
- Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. 1986, ApJ, 304, 15
- Baugh, C. M., Cole, S., Frenk, C. S., & Benson, A. J. 1999, MNRAS, 305, L21
- Benson, A. J., Cole, S., Frenk, C. S., Baugh, M., & Lacey C. G. 1998, preprint (astro-ph/9809171)
- Bernardeau, F. 1994, A&A, 291, 697
- Bernardeau, F. & Kofman, L. 1995, ApJ, 443, 479
- de Bernardis, P. *et al.* 1999, New Astronomy Reviews, 43, 289
- Bertschinger, E. & Gelb, G. 1991, Comput. Phys. 5, 164
- Blanton, M., Cen, R. Y., Ostriker, J. P., & Strauss, M. A. 1999, ApJ, 522, 590
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
- Borgani, S., Plionis, M., Coles, P., & Moscardini, L. 1995, MNRAS, 277, 1191
- Branchini, E., *et al.* 1999, MNRAS, 308, 1
- Bridle, S. L., Eke, V. R., Lahav, O., Lasenby, A. N., Hobson, M. P., Cole, S., Frenk, C. S., & Henry, J. P. 1999, MNRAS, 310, 565

- Cen, R. Y., & Ostriker, J. P. 1992, *ApJ*, 399, L113
- Cole, S., Weinberg, D. H., Frenk, C. S., & Ratra, B. 1997, *MNRAS*, 289, 37
- Coles, P., & Jones, B. 1991, *MNRAS*, 248, 1
- Colless, M. 1999, in *Proceedings of the ESO/Australia Workshop*, ed. R. Morganti, & J. C. Warrick, (Springer-Verlag), 9
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, 292, 371
- Davis, M. & Faber S.M. 1999, in *14th IAP Colloquium “Wide Field Surveys in Cosmology”*, ed. S. Colombi, Y. Mellier, & B. Raban (Gif-sur-Yvette: Editions Frontieres), 333
- Dekel, A. 2000, in *Cosmic Flows: Towards an understanding of Large Scale Structure*, ed. S. Courteau, M. Strauss, & J. A. Willick
- Dekel, A., Bertschinger, E., & Faber, S. M. 1990, *ApJ*, 364, 349
- Dekel, A., Eldar, A., Kolatt, T., Willick, J. A., Yahil, A., Faber, S. M., Courteau, S., & Burstein, D. 1999, *ApJ*, 522, 1
- Dekel, A., & Lahav, O. 1999, *ApJ*, 520, 24
- Dekel, A., & Rees, M. J. 1987, *Nature*, 326, 455
- Dekel, A., & Rees, M. J. 1994, *ApJ*, 422, L1
- Dressler, A. 1980, *ApJ*, 236, 351
- Eke, V. R., Cole, S., Frenk, C. S., & Patrick H. J. 1998, *MNRAS*, 298, 1145
- Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., & Schlegel, D. 1995, *ApJS*, 100, 69
- Freudling, W., *et al.* 1999, *ApJ*, 523, 1
- Guzzo, L., Strauss, M. A., Fisher, K. B., Gioveanelli, R., & Haynes, M. P. 1997, *ApJ*, 489, 37
- Hermit, S., Santiago, B. X., Lahav, O., Strauss, M. A., Davis, M., Dressler, A., & Huchra, J. P. 1996, *MNRAS*, 283, 709
- Jenkins, A., *et al.* 1998, *ApJ*, 499, 20
- Jing, Y. P. 1998, *ApJ*, 503, L9
- Jing, Y. P., & Suto, Y. 1998, *ApJ*, 494, L5
- Kaiser, N. 1986, *MNRAS*, 222, 323
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, *MNRAS*, 286, 795
- Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, *MNRAS*, 303, 188
- Kofman, L., Bertschinger, E., Gelb, M. J., Nusser, A., & Dekel, A. 1994, *ApJ*, 420, 44
- Kolatt, T., Dekel, A., Ganon, G., & Willick J. A. 1996, *ApJ*, 458, 419

- Lahav, O., Nemiroff, R. J., & Piran, T. 1990, ApJ, 350, 119
- Lahav, O., Rees, M., Lilje, P., & Primack, J. R. 1991, MNRAS, 251, 128
- Loveday, J., *et al.* 1995, ApJ, 442, 457
- Loveday, J., *et al.* 1999, in *14th IAP Colloquium “Wide Field Surveys in Cosmology”*, eds. S. Colombi, Y. Mellier & B. Raban (Gif-sur-Yvette: Editions Frontieres), 317
- Matarrese, S., Verde, L., & Heavens, A. 1997, MNRAS 270, 651
- Mo, H. & White, S. D. M. 1996, MNRAS, 282, 347
- Narayanan, V. K. & Weinberg, D. H. 1998, ApJ, 508, 440
- Perlmutter, S., *et al.* 1999, ApJ, 517, 565
- Porciani, C., Catelan, P., & Lacey C. 1999, ApJ, 513, L99
- Primack, J.R. 1998, in *Proceedings of the NATO Advanced Study Institute on Current Topics in Astrofundamental Physics*, ed. N. Sanchez, & A. Zichichi (Kluwer Academic Publishers), 469
- Riess, A. G., *et al.* 1998, AJ, 116, 1009
- Santiago, B. X., & Strauss, M. A. 1992, ApJ, 387, 9
- Saunders, W., *et al.* 2000, MNRAS, submitted, (astro-ph/0001117)
- Sheth, R., & Tormen, G. 1999, MNRAS, 308, 119
- Sigad, Y., Eldar, A., Dekel, A., Strauss, M. A., & Yahil, A. 1998, ApJ, 495, 516
- Somerville, R. S., Lemson, G., Sigad, Y., Dekel, A., Colberg, J., Kauffmann, G., & White, S. D. M. 2000, MNRAS, submitted, (astro-ph/9912073)
- Somerville, R. S., & Primack, J. R. 2000, MNRAS, in press (astro-ph/9806228)
- Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M. 1998, ApJ, 492, 428
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996 ApJ, 462, L17
- Strauss, M. A., Yahil, A., Davis, M., Huchra, J. P., & Fisher, K. B. 1992, ApJ, 397, 395
- Szapudi, I. 1998, ApJ, 497, 16
- Szapudi, I., & Colombi, S. 1996, ApJ, 470, 131
- Tonry, J. L., Blakeslee, J. P., Ajhar, E. A., & Dressler, A. 1997, ApJ, 475, 399
- Verde, L., Heavens, A., Matarrese, S., & Moscardini, L. 1998, MNRAS, 300, 747
- Wechsler, R. H., Gross, M. A. K., Primack, J. R., Blumenthal, G. R., & Dekel, A. 1998, ApJ, 506, 19
- Yahil A., Strauss M. A., Davis M., & Huchra J. P., 1991, ApJ, 372, 3

Zaroubi, S., Zehavi, I., Dekel, A., Hoffman, Y., & Kolatt, T. 1997, ApJ, 486, 21

**Table 1:** Recovery of the biasing function from the CDFs

$\Lambda$ CDM									
	halos vs. mass			galaxies vs. mass			early vs. late type		
	$z=0$	$z=1$	$z=3$	$z=0$	$z=1$	$z=3$	$z=0$	$z=1$	$z=3$
$\hat{b}$	0.67	1.21	2.98	0.89	1.31	2.38	1.11	1.32	1.28
$\hat{b}_a$	0.58	1.25	2.86	0.80	1.32	2.25	1.20	1.38	1.49
$\tilde{b}$	0.74	1.24	3.04	0.91	1.31	2.40	1.13	1.34	1.30
$\tilde{b}_a$	0.75	1.31	3.08	0.90	1.36	2.38	1.35	1.52	1.64
$\Delta$	0.16	0.14	0.11	0.08	0.08	0.08	0.55	0.38	0.56
$\tau$ CDM									
	halos vs. mass			galaxies vs. mass			early vs. late type		
	$z=0$	$z=1$	$z=3$	$z=0$	$z=1$	$z=3$	$z=0$	$z=1$	$z=3$
$\hat{b}$	0.90	2.18	6.62	0.93	1.71	4.44	1.17	1.34	1.27
$\hat{b}_a$	0.89	2.28	6.75	0.93	1.75	4.32	1.18	1.39	1.50
$\tilde{b}$	0.93	2.20	7.85	0.95	1.71	4.62	1.18	1.35	1.31
$\tilde{b}_a$	0.96	2.30	8.00	0.98	1.76	4.63	1.26	1.46	1.65
$\Delta$	0.08	0.07	0.20	0.08	0.04	0.08	0.22	0.20	0.54

**Table 2:** Redshift distortions and sampling errors in the biasing function

$\Lambda$ CDM							
	True	z-space	z-space ln	Volume	$l = 6^a$	$l = 8^a$	$l = 10^a$
$\hat{b}$	1.13	$1.12 \pm 0.006$	$1.09 \pm 0.02$	$1.12 \pm 0.05$	$1.17 \pm 0.05$	$1.23 \pm 0.04$	$1.31 \pm 0.06$
$\tilde{b}$	1.14	$1.12 \pm 0.006$	$1.10 \pm 0.02$	$1.13 \pm 0.05$	$1.17 \pm 0.05$	$1.24 \pm 0.04$	$1.32 \pm 0.06$
$\Delta$	—	$0.001 \pm 0.001$	$0.002 \pm 0.001$	$0.005 \pm 0.006$	$0.006 \pm 0.006$	$0.016 \pm 0.010$	$0.049 \pm 0.028$
$\tau$ CDM							
	True	z-space	z-space ln	Volume	$l = 6^a$	$l = 8^a$	$l = 10^a$
$\hat{b}$	1.188	$1.18 \pm 0.002$	$1.11 \pm 0.02$	$1.21 \pm 0.06$	$1.35 \pm 0.07$	$1.55 \pm 0.07$	$1.80 \pm 0.07$
$\tilde{b}$	1.192	$1.18 \pm 0.002$	$1.11 \pm 0.02$	$1.21 \pm 0.06$	$1.36 \pm 0.07$	$1.55 \pm 0.07$	$1.81 \pm 0.07$
$\Delta$	—	$0.002 \pm 0.0003$	$0.016 \pm 0.011$	$0.072 \pm 0.063$	$0.177 \pm 0.178$	$0.563 \pm 0.368$	$1.505 \pm 0.564$

<sup>a</sup> in units of  $h^{-1}\text{Mpc}$